

Simultaneous determination of helical unwinding angles and intrinsic association constants in ligand–DNA complexes: The interaction between DNA and calichearubicin B

(DNA unwinding/drug binding constants/daunorubicin/calicheamicin/ethidium bromide)

STEVEN M. ZEMAN*, KRISTOPHER M. DEPEW†, SAMUEL J. DANISHEFSKY†, AND DONALD M. CROTHERS*‡

*Department of Chemistry, Yale University, 225 Prospect Street, New Haven, CT, 06511; and †Laboratory for Bioorganic Chemistry, Sloan Kettering Institute for Cancer Research, 1275 York Avenue, New York, NY, 10021

Contributed by Donald M. Crothers, January 15, 1998

ABSTRACT We present a helical unwinding assay for reversibly binding DNA ligands that uses closed circular DNA, topoisomerase I (Topo I), and two-dimensional agarose gel electrophoresis. Serially diluted Topo I relaxation reactions at constant DNA/ligand ratio are performed, and the resulting apparent unwinding of the closed circular DNA is used to calculate both ligand unwinding angle (ϕ) and intrinsic association constant (K_a). Mathematical treatment of apparent unwinding is formally analogous to that of apparent extinction coefficient data for optical binding titrations. Extrapolation to infinite DNA concentration yields the true unwinding angle of a given ligand and its association constant under Topo I relaxation conditions. Thus this assay delivers simultaneous structural and thermodynamic information describing the ligand–DNA complex. The utility of this assay has been demonstrated by using calichearubicin B (CRB), a synthetic hybrid molecule containing the anthraquinone chromophore of (DA) and the carbohydrate domain of calicheamicin γ_1 . The unwinding angle for CRB calculated by this method is $-5.3 \pm 0.5^\circ$. Its K_a value is $0.20 \times 10^6 \text{ M}^{-1}$. For comparison, the unwinding angles of ethidium bromide and DA have been independently calculated, and the results are in agreement with canonical values for these compounds. Although a stronger binder to selected sites, CRB is a less potent unwinder than its parent compound DA. The assay requires only small amounts of ligand and offers an attractive option for analysis of DNA binding by synthetic and natural compounds.

By now there is little doubt that intercalation is a phenomenon of widespread biological importance. The intercalation model (1) describes a lengthening of the DNA helix and a concomitant unwinding of the phosphodiester backbone to accommodate a molecule of intercalator. As this unwinding varies with intercalator, the unwinding angle (ϕ) is a valuable biophysical parameter when evaluating the ability of a compound to interact with DNA. It represents the number of degrees by which the DNA is unwound about its helical axis per ligand molecule bound.

To date, most experimental approaches for measuring ϕ have exploited the special conformational characteristics of closed circular DNA (ccDNA). As isolated from prokaryotes and viruses, ccDNA is negatively supercoiled, that is to say, its writhe (Wr), linking difference (ΔLk), and superhelical density (σ , the average number of superhelical turns per 10 bp) are negative. § Upon binding by intercalator, ccDNA is unwound so that ΔLk and σ each become less negative (Fig. 1B). For a certain critical amount of bound intercalator, all superhelicity

disappears and the ccDNA exists as a relaxed circle in which the helical axis lies flat in a plane. For this example, ΔLk , Wr , and σ equal zero (Fig. 1C). Further titration with intercalator past this topological equivalence point continues to unwind the helix, introducing supercoils in an opposite sense so that both ΔLk and σ become positive (Fig. 1D).

Early hydrodynamic methods of studying σ and ϕ were based on changes in the sedimentation velocity (4–8), buoyant density (9–11), and viscosity (8) of ccDNA when bound by intercalator. The discovery of the enzyme topoisomerase I (Topo I) (12, 13) permitted other experimental approaches to both ϕ and σ , which were extended to studies of complexes between DNA and such topology-perturbing proteins as *EcoRI* (14), *Xenopus* transcription factor A (15), RNA polymerase (16–18), and the *lac* repressor (19). Topo I facilitates the thermodynamic equilibration between supercoiled and relaxed forms of ccDNA by creating a transient single-stranded nick at which superhelical strain can be relieved. Methods for the determination of ϕ and σ that have used Topo I and intercalating ligand are based on the processes shown in Fig. 1.

When ccDNA that has been intercalated to some extent is acted on by Topo I, a fully relaxed species is generated whose twist (Tw) and Wr reflect the presence of bound intercalator. In this fully relaxed species, $Wr = \Delta Lk = 0$. When this intercalator is removed, the sudden change in Tw is partially offset by a compensatory Wr so that Lk , an invariant topological property intrinsic to any intact molecule of ccDNA, remains unchanged. This change in Wr manifests itself conformationally as a reappearance of superhelicity, the magnitude of which is determined by the original amount of intercalator bound at the time of enzymatic religation. Quantitating this superhelicity in some way enables a calculation of ϕ or σ . In reality, the resulting ccDNA is not homogeneous and actually exists as a family of topoisomers differing by integral increments of ΔLk . Thermal conformational fluctuations at the time of ccDNA religation give rise to a Boltzmann distribution of topoisomers around an average ΔLk (20–23). This central ΔLk (ΔLk_c) is taken as the average most abundant

Abbreviations: ccDNA, closed circular DNA; Topo I, topoisomerase I; CRB, calichearubicin B; DA, daunorubicin; CLM, calicheamicin γ_1 ; EB, ethidium bromide; Wr , writhe; ΔLk , linking difference; Tw , twist; 1D, one dimensional; MGOS, methyl glycoside of the CLM oligosaccharide.

A commentary on this article begins on page 4092.

‡To whom reprint requests should be addressed.

§ $Lk = Tw + Wr$, where Tw is the twist about the helical axis and Wr is the writhe of the helical axis through three-dimensional space. Whereas Tw and Wr are each geometric properties that vary with environmental conditions such as temperature and ionic strength, Lk is a topological property that is invariant in the absence of single-stranded nicks. ΔLk is defined as the difference in superhelical turns between a fully relaxed species for which $Wr = 0$ and species j of nonzero Wr . Although ΔLk can assume a fractional value, Lk is integral by definition. Excellent treatments of ccDNA topology exist in the literature (2, 3).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/954327-6\$2.00/0
PNAS is available online at <http://www.pnas.org>.

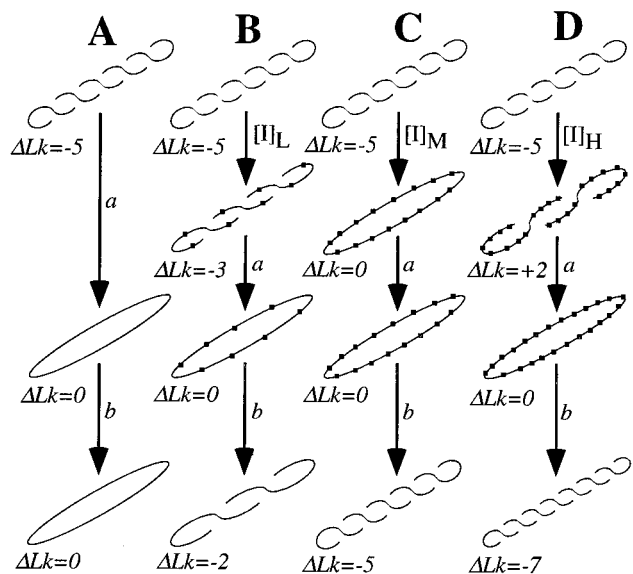


FIG. 1. Schematic Topo I relaxation reactions of ccDNA at no, low ($[I]_L$), medium ($[I]_M$), and high ($[I]_H$) ligand binding. *a*, Topo I relaxation; *b*, ligand removal by organic extraction. Ligand molecules are represented as nodules along the helical axis of the ccDNA shown. Reaction **A** lacks ligand, so the ccDNA product is fully relaxed with $\Delta Lk = 0$. In reaction **B**, enough ligand is bound to titrate out 2 superhelical turns (sht) so that the ccDNA now contains -3 sht. After *a* and *b*, these 2 titrated sht are manifested in a final ΔLk of -2 . Reaction **C** depicts just enough bound ligand to fully titrate all 5 sht, so product and starting ccDNAs are indistinguishable. Finally, **D** shows high-ligand binding to $+2$ sht beyond the equivalence point of **C**. In this example, the ccDNA product bears this additional superhelicity in its ΔLk of -7 .

topoisomer in the Boltzman population and is used as a comparative value between populations arising from Topo I reaction on ccDNA with various intercalator concentrations.

Probably the simplest method to quantitate such topoisomeric reaction products is resolution by agarose gel electrophoresis. In this method, the speed of topoisomer migration through a gel is directly proportional to degree of superhelicity. Electrophoretic band counting methods (20, 21, 24) differ from earlier procedures in that they eliminate the requirement for costly equipment such as the ultracentrifuge and are in general easier to perform. However, because band counting methods resolve topoisomeric reaction products in only one dimension, they suffer from severely limited resolution at the topological extremes of either fully relaxed or highly supercoiled ccDNA.

Unwinding as a Viable Measure of Extent of Ligand Binding. The present method to calculate ϕ relies on resolving topoisomeric reaction products in two electrophoretic dimensions (25, 26) and analyzing the degrees of unwinding in a way that is analogous to the classic treatment of optical binding data (27). The incorporation of a second gel dimension, run perpendicular to the first, allows the clear resolution of topoisomers, which otherwise comigrate in just one dimension.

Mathematical analysis of reaction products uses the Scatchard binding isotherm adjusted for nearest neighbor exclusion by bound ligand (28, 29), presented in closed form by McGhee and Von Hippel (30):

$$\frac{r}{C_F} = K_a(1 - nr) \left[\frac{1 - nr}{1 - (n-1)r} \right]^{n-1}, \quad [1]$$

where C_F is the free ligand concentration, r is the ratio of bound ligand to base pairs, K_a is the intrinsic association constant under the chosen conditions, and n is the number of sites (normally base pairs) covered by one molecule of bound

ligand. Eq. 1 simplifies to linear form at low binding when $r \ll 1/n$:

$$r/C_F = K_a[1 - (2n - 1)r]. \quad [2]$$

Traditionally, r has been calculated optically via the measured apparent ligand concentration coefficient difference between free and bound states, $\Delta\epsilon_{ap}$. With increasing DNA concentration, $\Delta\epsilon_{ap}$ approaches $\Delta\epsilon$, and the fraction of bound ligand is $\Delta\epsilon_{ap}/\Delta\epsilon$. With $C_B = (\Delta\epsilon_{ap}/\Delta\epsilon)C_T$,

$$r = \frac{C_B}{C_N} = \frac{\Delta\epsilon_{ap}C_T}{\Delta\epsilon C_N}, \quad [3]$$

where C_B , C_T , and C_N are, respectively, the concentrations of bound ligand, total ligand, and total DNA. With increasing dilution $C_F \approx C_T$, so substitution of Eq. 3 into Eq. 2 and expansion to first order in $(\Delta\epsilon - \Delta\epsilon_{ap})$ yields the linear equation:

$$\Delta\epsilon_{ap} = \Delta\epsilon - \frac{\Delta\epsilon_{ap}}{K_a[C_N - (2n - 1)C_T]}. \quad [4]$$

A plot of $\Delta\epsilon_{ap}$ vs. $\Delta\epsilon_{ap}/[C_N - (2n - 1)C_T]$ intercepts the ordinate at $\Delta\epsilon$ as C_N approaches infinity. The slope of the extrapolation yields $-1/K_a$ directly.

It is Eq. 4 that has been brought to bear upon the problem of unwinding angle determination. For this task, $\Delta\epsilon_{ap}$ and $\Delta\epsilon$ are replaced by, respectively, ϕ_{ap} and ϕ , the apparent unwinding elicited by incomplete binding of ligand where $r \ll 1/n$ and the unwinding angle of the ligand itself at extrapolated infinite C_N . Hence, Eq. 4 can be rewritten to describe the binding equilibrium of ligand to DNA by monitoring apparent unwinding:

$$\phi_{ap} = \phi - \frac{\phi_{ap}}{K_a[C_N - (2n - 1)C_T]}. \quad [5]$$

The resulting line's interception with the ordinate yields ϕ and the negative inverse of its slope gives K_a under the temperature and ionic conditions of the Topo I relaxation reaction.

Topo I relaxation reactions containing identical total amounts of DNA and ligand are run in a series of ascending volumes. As total reaction volume increases, dissociation of the ligand-DNA complex is promoted by the parallel decrease of both DNA and ligand concentrations. The bp/ligand ratio remains unchanged (and high) throughout, ensuring the low r required by the mathematical approximations above. Reactions are treated and resolved as described below to determine ΔLk_c , the Boltzman center of mass of integrated topoisomer band intensity representing the average most populous topoisomer (Fig. 2). For a given distribution of topoisomeric reaction products, ϕ_{ap} is determined as follows:

$$\phi_{ap} = \frac{360N_D(\Delta Lk_c - \Delta Lk_c^0)}{N_L}, \quad [6]$$

where ΔLk_c and ΔLk_c^0 are the Boltzman centers of topoisomer band intensity of, respectively, a reaction with ligand and a control reaction containing no ligand (this difference is $\Delta\Delta Lk_c$). N_D and N_L represent, respectively, the actual numbers of ccDNA and ligand molecules in a particular reaction. The ligand is presumed to be randomly distributed throughout all the ccDNA molecules present in the reaction mixture. The factor of 360 represents 360° per ΔLk in the ccDNA, so ϕ_{ap} carries the expected units of degrees. With ϕ_{ap} , Eq. 5 can be used to obtain ϕ and K_a . The required estimation of n is discussed below.

Intercalation Directed by a Selective Minor Groove Binder: Calichearubicin B (CRB). In the intercalation of DNA by the antineoplastic agent daunorubicin (DA), the daunosamine sugar is indispensable (Fig. 4). Although only the aromatic

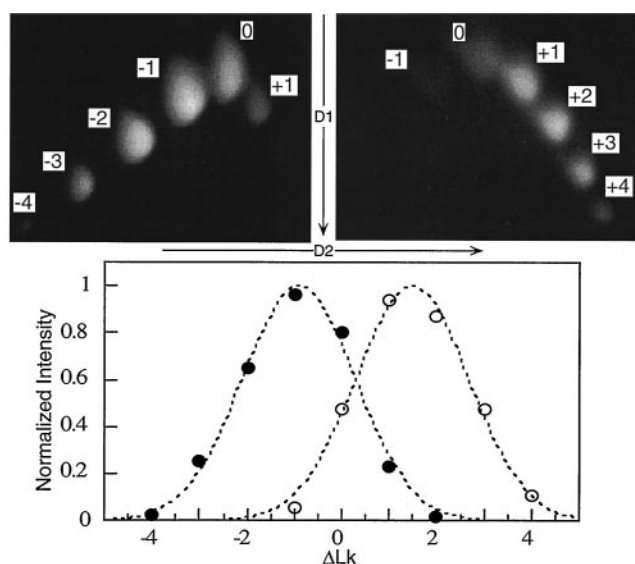


FIG. 2. Example calculation of ϕ_{ap} . Photos of EB-stained gels are shown above with their corresponding Boltzman distributions below. The results shown are from reactions containing (*Left*) (●) and lacking (*Right*) (○) the unwinding ligand DAUN, and values for ΔLk as well as electrophoretic dimensions are indicated on the gels. ΔLk_c is -0.951 for the ligand-containing reaction, and the maximum of the ligand-free control reaction, ΔLk_c^0 , is $+1.49$. This implies a $\Delta\Delta Lk_c$ of -2.44 . With $N_L = 1.46 \times 10^{14}$ ligand per reaction and $N_D = 1.17 \times 10^{12}$ ccDNA per reaction, Eq. 6 gives $\phi_{ap} = -7.03^\circ$. This becomes one point in a binding isotherm as shown in Fig. 3.

rings of the DA anthraquinone are directly intercalated, it is the daunosamine sugar that rests in the minor groove to electrostatically stabilize the drug–DNA complex (31, 32). Indeed, in the absence of this sugar, daunorubicinone does not exhibit any appreciable intercalative propensity at physiologically relevant ionic strength and pH (data not shown). Yet for all its stability, the DA–DNA interaction demonstrates only minimal sequence specificity (33, 34).

In contrast, the enediyne antibiotic calicheamicin γ_1^I (CLM) exhibits marked sequence specificity in its cleavage of double-stranded DNA (36, 37). In this reaction, it is the extended pentacyclic carbohydrate domain of CLM that recognizes and directs aglycone cleavage to polypyrimidine–polypurine (Py·Pu) tracts in DNA (38, 39). This sequence selectivity has been proposed to result from a number of factors, including specific van der Waals contacts between atoms on the CLM carbohydrate and the DNA (40), preorganization of the CLM carbohydrate into an extended rigid structure (41, 42), and a greater inherent propensity for distortion of local DNA topology at preferred sequences, which promotes binding by CLM (43, 44). Of these, the latter two have emerged as the most critical for sequence selective binding by CLM. The specificity observed seems to be determined more by general DNA shape (minor groove width and depth, increased helical repeat, and bending) than by distinct contacts between CLM and DNA. With their deepened and narrowed minor groove and their slightly increased helical repeat, Py·Pu tracts present CLM with an environment spatially amenable to binding by the extended carbohydrate (45, 46).

With these considerations in mind, we wondered whether a synthetic union of the CLM carbohydrate with the DA anthraquinone might result in a molecule in which the binding specificity of the former is conferred to the intercalation of the latter. Such a molecule is CRB, shown in Fig. 4. The two domains of CRB are joined by an ethoxyethyl spacer, judged by computer modeling studies as sufficient to accommodate both anthraquinone intercalation and sensing of minor groove environment by the carbohydrate. Preliminary studies of the binding interaction between CRB and DNA indicated that the

molecule binds preferentially to Py·Pu tracts (47), suggesting that the known sequence selectivity of the carbohydrate domain is in fact conferred to the hybrid molecule. Furthermore, the absorption profile obtained upon binding CRB to bulk DNA correlated essentially exactly in intensity attenuation and shift of λ_{max} to that of natural DA, suggesting that the CRB chromophore is fully intercalated (47).

These findings spurred the attempt to quantify this binding behavior by the present helical unwinding assay. However, the mathematical treatment of this binding had to reflect the fact that CRB does not bind to bulk DNA in the nonspecific manner shown by other compounds such as DA and ethidium bromide (EB). For binding by CRB, the number of potential binding sites was calculated from the known sequence of the ccDNA used in the helical unwinding assay. In the CRB studies, C_N is therefore the concentration of potential binding sites rather than of base pairs.

MATERIALS AND METHODS

Materials. Purified pUC19 DNA was isolated from transformed JM101 *Escherichia coli* by the Qiagen protocol. Assay by one-dimensional (1D) agarose gel electrophoresis confirmed that the DNA was more than 98% supercoiled ccDNA. Calf thymus Topo I and ultrapure agarose were purchased from GIBCO/BRL. EB and DA were purchased from Aldrich Chemicals and used without further purification. Microcon centrifugal concentrators were purchased from Amicon. All water used in reactions, gel casting, and running buffers was double-distilled.

Reactions. Before each reaction series, ligand concentration as determined by absorbance by using the following molar extinction coefficients: $\epsilon_{480} = 5,600 \text{ M}^{-1}\cdot\text{cm}^{-1}$ for EB and $\epsilon_{480} = 11,500 \text{ M}^{-1}\cdot\text{cm}^{-1}$ for DA and CRB. DNA was prepared from stock solutions as a single premixture containing appropriate buffer for the smallest-volume reaction. This premixture was then distributed equally into the reaction tubes. Ample excess premixture was prepared so that the base-pair concentration could be determined exactly by absorbance. For this, a molar extinction coefficient of $\epsilon_{260} = 13,200 \text{ M}(\text{bp})^{-1}\cdot\text{cm}^{-1}$ was used.

Reactions containing identical amounts of ccDNA and ligand were run in a series of volumes spanning two orders of magnitude (20 μl to 2 ml) in intervals of factors of two. DNA concentration was typically between 175 and 275 μM (bp) in the smallest-volume reaction. Ligand concentration depended on the potency of ligand unwinding (see *Discussion*) but was typically between 5 and 10 μM in the smallest-volume reaction. Reactions of higher volume were supplemented with appropriate additional buffer stock to ensure constant ionic strength in all reactions.

Unwinding reactions were run in 50 mM Tris-HCl, pH 7.5/50 mM KCl/10 mM MgCl_2 /0.5 mM DTT/100 μM EDTA/BSA (30 $\mu\text{g}/\text{ml}$). Complete relaxation of the ccDNA was achieved using 10–11 units of Topo I per μg of ccDNA in each reaction and was ensured by reacting an additional control sample containing ccDNA but no ligand. After reaction for 4 h at 37°C, samples were extracted with an equal volume of 1:1 PhOH/ CHCl_3 buffered to pH 7.5 with 50 mM Tris-HCl to remove protein and ligand. The phases were partitioned and the aqueous phase from each sample was concentrated to approximately 10 μl by centrifugation. Small (about 0.5 μl) aliquots were removed for preliminary analysis by 1% 1D agarose gel electrophoresis. Control reactions lacking unwinding ligand were identically reacted, extracted, and concentrated as those containing ligand.

Nonuniform binding of EB during gel staining might be expected to introduce artifacts in observed band intensity that could skew the calculated value of ΔLk_c . Both the superhelical state of the ccDNA and the absolute amount of ccDNA in a given band might be expected to cause such nonuniformity in EB binding. Two sets of control experiments were performed to investigate these possibilities. In one, identical amounts of ccDNA intercalated by various amounts of EB were com-

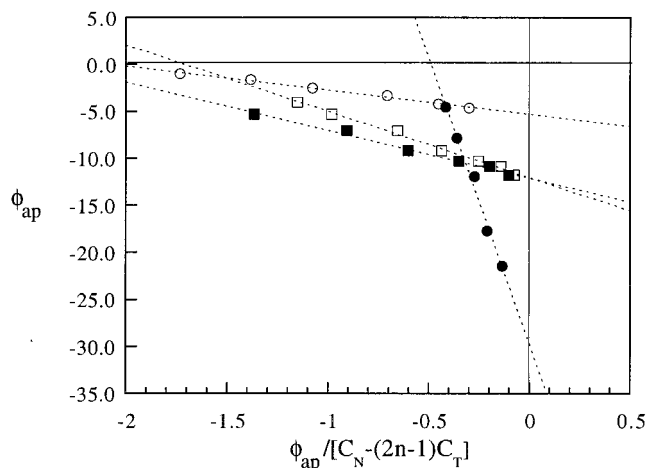


FIG. 3. Binding isotherms for CRB (\circ , $n = 1$), DA (\square , $n = 4$), and EB (\bullet , $n = 2$). Unwinding data are plotted according to Eq. 5. Interception of the ordinate at extrapolated infinite C_N yields ϕ directly, and the negative inverse of the slope yields K_a . An additional hypothetical isotherm for DA (\blacksquare , $n = 6$) has been included to show that the intercept value of ϕ does not depend on the value of n used.

pletely relaxed and treated as described above so that the product ccDNAs spanned a range of superhellicities. These products were resolved by 1D agarose gel electrophoresis and visualized by EB staining. For the range of superhellicities relevant for the present method, the observed intensity (quantitated as described below) showed no systematic dependence on ΔLk (data not shown). In another control, various absolute amounts of ccDNA homogeneous in ΔLk were resolved by 1D agarose gel electrophoresis and visualized by EB staining. Quantitation of total topoisomer band intensity (see below) showed a linear dependence of intensity on the amount of ccDNA up to 1.2 μg of ccDNA, after which point the intensity saturated (data not shown). Because the present method involves no two-dimensional topoisomer band containing more than 0.6 μg of ccDNA, all quantitations of band intensity are well within the linear range required for reliable quantitation.

Electrophoresis. Two-dimensional gel casting units (each 17×17 cm) were constructed according to established protocols (25). Gels were cast with 250 ml of 1% agarose in 80 mM TBE (80 mM Tris borate/1 mM EDTA) to a final gel thickness of approximately 8 mm. A single loading well with a diameter of 1 mm was made in a corner of each gel. The first dimension was electrophoresed 20 h in 1 liter of 80 mM TBE running buffer at 1.9 V/cm. After the first dimension, the gel was rotated 90° in its casting tray, chloroquine was added in a small volume to a final concentration of 1.1 $\mu\text{g}/\text{ml}$ and the gel was soaked with light agitation for 1 h. Electrophoresis in the second dimension then proceeded at 1.9 V/cm for an additional 18 h.

Visualization and Analysis. After electrophoresis, gels were soaked in EB (1 $\mu\text{g}/\text{ml}$) for 2 h, destained in water for 1 h, and photographed on a UV transilluminator with Polaroid 667 film. The photographs were scanned with a Hewlett-Packard Scan Jet 3P and the photonegatives of individual topoisomer bands were quantitated by using NIH IMAGE software. Resulting datasets were analyzed by using KALEIDAGRAPH software, with an appropriate baseline being subtracted from each plot prior to integration of total topoisomer band intensity. The integrated intensities were then plotted to give a Boltzman distribution from which ΔLk_c was calculated by fitting to $I = I_M \exp[-w(\Delta Lk - \Delta Lk_c)]$, where I and I_M are, respectively, integrated band intensity and maximum band intensity at the peak of the topoisomer distribution and w is the width of the distribution. ΔLk_c was then used to obtain ϕ_{ap} via Eq. 6 (see

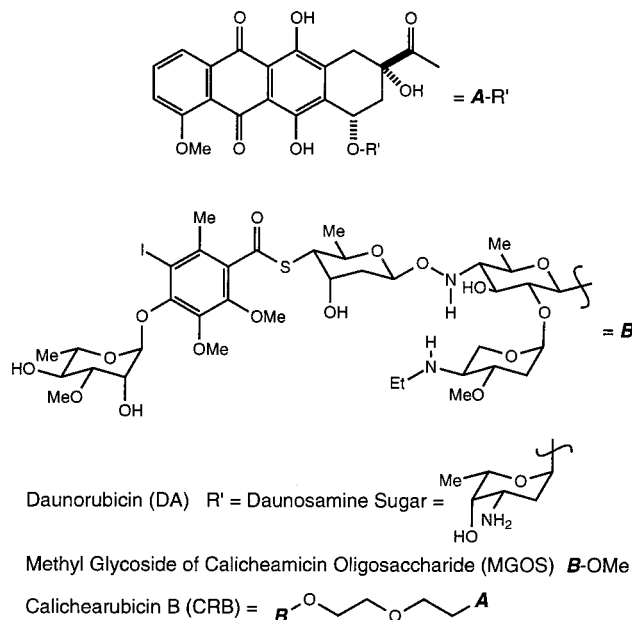


FIG. 4. Structures of DA, MGOS, and CRB. Note that the connectivity in CRB is through the β -anomer of B .

above). Fig. 2 shows an example of topoisomeric reaction products from reactions with and without ligand, and the corresponding Boltzman distributions of topoisomer band intensity.

RESULTS

EB and DA as Comparative Unwinders. To demonstrate the accuracy of the method and provide comparative values for CRB, helical unwinding assays were performed on the known intercalators DA and EB. The results of these reactions are displayed graphically in Fig. 3. The values of n used were 2 for EB and 4 for DA (see *Discussion*).

Fig. 3 shows that the line for EB intercepts the ordinate at $-28.7 \pm 2.9^\circ$ and the line for DA intercepts at $-12.1 \pm 1.2^\circ$, both values in good agreement with canonical values for the unwinding angles (Table 1). The value for ΔLk_c^0 , the Gaussian center of topoisomeric mass for a control reaction lacking any unwinding ligand, was found to be $+1.5^\circ$ for the ccDNA used. This value (discussed below) is an average of 5 ligand-free controls and is used as ΔLk_c^0 in all calculations of $\Delta \Delta Lk_c$.

Unwinding by CRB. Methidiumpropyl EDTA-Fe(II) footprinting data for CRB and the methyl glycoside of the CLM oligosaccharide (MGOS) on a *SphI*-*SgrAI* restriction fragment of pBR322 showed cleavage enhancement at identical sites for CRB and MGOS, each corresponding to Py-Pu sites of 4 bp. The observed sites d(CTTC)·d(AGGA) and d(TTTC)·d(GAAA) extend the list of other polypyrimidine sites observed for MGOS and CLM in the literature, suggesting that it is indeed more the general steric environment of the minor groove in such

Table 1. Summary of unwinding and binding parameters for CRB, DA, and EB

Ligand	ϕ	ϕ_c	K_a , $\times 10^6 \text{ M}^{-1}$
CRB	-5.30°	—	0.39
DA	-12.1°	-12°	0.14
EB	-28.7°	-26°	0.02

ϕ is the unwinding angle calculated by the present method. ϕ_c is the canonical unwinding angle for a given ligand, obtained as the intercept of a respective isotherm with the ordinate in Fig. 3. K_a is the intrinsic association constant calculated as the negative inverse of a respective isotherm's slope in Fig. 3.

motifs that facilitates carbohydrate binding rather than distinct ligand–DNA contacts of certain unique sequences. With this in mind, 380 potential binding sites were calculated in pUC19 (the ccDNA used in the experiments) by taking all possible 4-bp combinations of cytidine and thymidine (or guanosine and adenine for the complementary strand). Here C_N is taken as the concentration of such sites in an unwinding reaction and $n = 1$. Fig. 3 shows the graphical representation of these data, giving an unwinding angle of $-5.3 \pm 0.5^\circ$. The association constants for CRB, DA, and EB obtained by taking the negative inverse of the slopes in Fig. 3 are summarized in Table 1.

DISCUSSION

Comparing the Binding of CRB and DA. The higher K_a of CRB relative to its parent intercalator DA would seem to indicate an enhanced thermodynamic complex stabilization gained by replacing the daunosamine sugar with the tethered CLM carbohydrate. However, the attenuated ϕ of CRB ($-5.3 \pm 0.5^\circ$) relative to DA ($-12.1 \pm 1.2^\circ$) suggests that other factors are also at play.

One interpretation is that the unwinding effect of the chromophore in CRB is partially canceled by a positive winding by the bound carbohydrate at its minor groove site. Such partial cancellation would diminish apparent CRB unwinding relative to DA. In fact, the carbohydrate domain has been reported to bend DNA in NMR solution studies (41, 46). Although Topo I experiments using MGOS alone indicate a positive $\Delta\Delta Lk$ (data not shown), this effect is only observable with levels of MGOS in excess of actual binding sites (saturating conditions such as these invalidate the equations of the assay; see below), and even then, the effect is only slight ($\Delta\Delta Lk \approx +0.75$). It is therefore implausible that the much lower levels of this weak winder in CRB would be capable of compensating the relatively potent unwinding effect of the anthraquinone by more than half. The observation of a slightly positive $\Delta\Delta Lk$ for MGOS confirms earlier hypotheses that methidiumpropyl/EDTA Fe(II) cleavage enhancement at MGOS-bound sites derives from local topological perturbation of DNA (39, 47). However, the present assay is incapable of distinguishing ΔTw from ΔWr in the global ΔLk actually observed. It thus remains unknown whether this slight winding effect by MGOS is because of twisting, writhing, or a combination of both.

Mathematical analysis of the CRB data identical to that used for DA, using Eq. 5 with $C_N = [bp]_{TOT}$ and $n = 4$ indicates a 3-fold higher total binding affinity of DA relative to CRB. This treatment ignores any inherent sequence preference of CRB and serves only to compare the apparent association constants of the two drugs for DNA of average sequence. Because the apparent affinity analyzed this way is smaller for CRB than for DA, we conclude that the sequence selectivity of the CLM carbohydrate serves to restrict the binding of CRB from sites otherwise available to the parent intercalator DA.

The selectivity of CRB in its binding of DNA is supported by methidiumpropyl/EDTA·Fe(II) footprinting studies in which patterns of cleavage enhancement for CRB and MGOS correlate exactly. Adjusting C_N to reflect the consensus Py·Pu motif required for CRB binding and setting $n = 1$ accounts for CRB binding selectivity in Eq. 5 and raises the K_A value for CRB above that for DA. However, the K_a values in Table 1, which are based on this reasoning, imply only an approximately 3-fold improvement of CRB affinity for sites of appropriate sequence over the affinity of DA for bulk or average-sequence DNA. Thus the carbohydrate side chain provides only a modest advantage over the daunosamine sugar in contributing to affinity for the target sequences. The major effect of exchanging the carbohydrate portion is to prevent binding of CRB to the bulk sequences to which DA can readily bind. This result is not entirely expected, given the high DNA binding affinity of MGOS that can be inferred from its ability to inhibit DNA cutting by CLM at MGOS

concentrations only an order of magnitude higher than CLM (39).

Although the synthetic union of the DA chromophore and the CLM carbohydrate has resulted in a molecule combining minor groove sequence recognition with intercalative unwinding, this has been at the cost of unwinding propensity. The data support a model in which the chromophore- and carbohydrate-binding domains in the DNA twist toward each other by $+6$ to 7° so that simultaneous binding by each drug domain is possible. In this model, the -12° unwinding angle of the DA chromophore is partially canceled in its contribution to the global parameter ΔLk so that the net ϕ observed is only -5.3° . Such a scenario encompasses earlier spectrophotometric evidence supporting full intercalation by the CRB chromophore by comparison to that of DA (47), as well as the finding that MGOS alone does not engender appreciable positive twist in the DNA at its binding site (see above). These findings suggest that the ethoxyethyl linker in CRB is of a slightly nonoptimum length to allow simultaneous bifunctional drug–DNA interactions without a deformation of the DNA helix in the form of positive twisting. This is not unreasonable given the unavoidable constraint to segments of integral bond length in the synthesis of the spacer separating the two domains of CRB. This postulated imperfect steric match may be the reason why the carbohydrate side chain does not contribute more strongly than it does to the affinity of CRB for its target sequences.

The Mutual Dependence of ϕ and σ . A difficulty with past methods for determination of ϕ involves the mutual dependence of ϕ and σ . Traditionally, determination of one of these two parameters required prior knowledge of the other, precluding any *a priori* calculation of either. In addition, it has become common practice in most of the past methods to calibrate the experimental ϕ of an unknown ligand to that of EB, itself accepted to be about $-26 \pm 2^\circ$ (9). Hence any calculated unwinding angle was necessarily a relative one. Although initial electrophoretic band counting was in fact capable of providing an absolute measure of ϕ , the method still presupposed prior knowledge of σ for the ccDNA used in the reactions and a measure of the maximum bound mole fraction of ligand per base pair (ν).

In contrast, the present assay for the simultaneous determination of ϕ and K_a requires no prior knowledge of either σ or ν , using instead sufficiently high C_N/C_T ratios so that extrapolation to infinite C_N is possible. At this infinity point, the ligand is assumed to be essentially completely bound to the DNA, and $\phi_{ap} = \phi$. The prerequisites for the present method are (i) that the K_d of the ligand–DNA complex be sufficiently low to warrant the assumption of near complete ligand binding at high C_N , (ii) knowledge of the size (in base pairs) of the ccDNA used in the relaxation reactions, and (iii) an estimate of the number of potential binding sites along a molecule of the ccDNA used.

Choice of n . Consideration of Eq. 5 reveals the need for an estimate of n , the number of potential binding sites covered by one molecule of bound ligand. For intercalating ligands such as DA and EB, which show only minimal sequence selectivity in their binding to bulk DNA, each base pair in the ccDNA may be seen as making an equal contribution to the number of total viable binding sites. Under these conditions, potential binding sites overlap one another and the neighbor exclusion model (28–30) applies, with n greater than unity [notice that substitution of $n = 1$ into Eq. 1 regenerates the original Scatchard binding isotherm (27) in which independent potential binding sites do not overlap]. If the serially diluted reactions are run at a sufficiently high DNA/ligand ratio, then $C_T \ll C_N$ and the choice of n in Eq. 5 has little effect on the resulting plot. For less potent unwinders, it may be necessary to run reactions of higher ligand concentration to generate a quantifiable $\Delta\Delta Lk_c$. In these situations, the C_T/C_N quotient rises and the choice of n becomes more important. However even in this case, n affects only the slope of the line in Eq. 5 (the source of K_a) and

has no bearing upon the line's intercept with the ordinate, from which the unwinding angle is determined (Fig. 3). Given a reasonable estimate of the number of potential binding sites along the ccDNA, this method is, therefore, capable of generating a good measure of ϕ regardless of the chosen n . For exceedingly weak unwinders, a detectable $\Delta\Delta Lk_c$ is sometimes possible only by addition of excess ligand so that all potential binding sites remain saturated even at low C_N . At this point, Eq. 5 becomes invalid. The limit of the method's applicability is thus reached as the C_T/C_N ratio approaches unity.

The choices of $n = 2$ and $n = 4$ for the respective ligands EB and DA are based on literature values for these compounds (48, 49). Because of DA's importance as a comparative model for anthracycline binding in the present study of CRB, its interaction with bulk DNA was carefully monitored in fluorescence quenching titrations under Topo I reaction buffer conditions. Analysis of fluorescence emission intensity by Eq. 1 corroborated the value of $n = 4$ used for DA in the unwinding calculation (data not shown). The value of $n = 1$ for the present CRB binding calculation reflects the fact that the number of base pairs in a potential CRB binding site was preset to 4 in the calculation of these sites on the ccDNA. Because C_N in this case represents the concentration of sites and one molecule of CRB is assumed to cover one site, n is set to unity. The mathematical treatment of CRB unwinding data is thus consistent with a classical Scatchard interpretation in which independent potential binding sites do not overlap. This is in fact almost entirely the case with the potential CRB binding sites in the ccDNA used.

Regarding ΔLk_c^0 and ΔLk_c . When relaxing ccDNA with Topo I in the absence of ligand, theory predicts that the average most populous topoisomer should center on the fully relaxed species for which $\Delta Lk_c^0 = 0$ (Fig. 1A). Yet the value for ΔLk_c^0 used in the above calculations is not zero, but +1.5. This shift of +1.5 superhelical turns away from the expected zero results from the fact that the enzymatic relaxation reaction is carried out at higher ionic strength than the subsequent two-dimensional agarose gel analysis. However, this has no effect on the calculation of $\Delta\Delta Lk_c$, because both ligand-free and ligand-containing samples are reacted and treated identically. Because superhelicity of ccDNA depends on ionic strength, it is of paramount importance to ensure identical salt conditions for all samples during reaction. To this end, supplementing samples of higher volume with appropriate additional reaction buffer is essential. It is then merely a matter of convention which topoisomer one chooses to use as a reference to compare control and ligand-containing samples. For this, the Boltzmann center of topoisomer band intensity is the most obvious and most accurately calculable point of reference for the determination of $\Delta\Delta Lk_c$.

In summary, the present method is unique among related assays for its efficiency in describing ligand-DNA interactions both structurally and thermodynamically with a single experiment. The assay's ease of performance and its requirement for only extremely small amounts of ligand make it ideal for the analysis of synthetic and natural compounds intended as potential DNA-interacting drugs. Disadvantages include the restriction to solution conditions suitable for Topo I activity and the requirement that the ligand produce a change in ΔLk so that its binding can be detected.

This work was supported by the National Institutes of Health (Grants 5-T32-AI 07404-05, CA 28824, and HL 28549).

- Lerman, L. S. (1961) *J. Mol. Biol.* **3**, 18–30.
- Bates, A. D. & Maxwell, A. (1993) in *DNA Topology*, eds., Rickwood, D. & Male, D. (Oxford Univ. Press, Oxford), pp. 17–61.
- Cozzarelli, N. R., Boles, T. C. & White, J. H. (1990) in *DNA Topology and Its Biological Effects*, eds. Cozzarelli, N. R. & Wang, W. C. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 139–184.
- Wang, J. C. (1969) *J. Mol. Biol.* **43**, 263–272.
- Bauer, W. & Vinograd, J. (1968) *J. Mol. Biol.* **33**, 141–171.
- Waring, M. (1970) *J. Mol. Biol.* **54**, 247–279.
- Crawford, L. V. & Waring, M. J. (1967) *J. Mol. Biol.* **25**, 23–30.
- Révet, B. M. J., Schmir, M. & Vinograd, J. (1971) *Nat. New Biol.* **229**, 10–13.
- Wang, J. C. (1974) *J. Mol. Biol.* **89**, 783–801.
- Gray, H. B., Jr., Upholt, W. B. & Vinograd, J. (1971) *J. Mol. Biol.* **62**, 1–19.
- Radloff, R., Bauer, W. & Vinograd, J. (1967) *Proc. Natl. Acad. Sci. USA* **57**, 1514–1521.
- Wang, J. C. & Liu, L. F. (1979) in *Molecular Genetics*, ed. Taylor, J. H. (Academic, New York), Part III, pp. 65–88.
- Wang, J. C. (1983) in *Nucleic Acid Research*, eds. Mizobuchi, K., Watanabe, I. & Watson, J. D. (Academic, New York), pp. 549–566.
- Kim, R., Modrich, P. & Kim, S.-H. (1984) *Nucleic Acids Res.* **12**, 7285–7292.
- Hanas, J. S., Bogenhagen, D. F. & Wu, C.-W. (1984) *Nucleic Acids Res.* **12**, 1265–1276.
- Beard, P., Hughes, M., Nyfeler, K. & Hoey, M. (1984) *Eur. J. Biochem.* **143**, 39–45.
- Amouyal, M. & Buc, H. (1987) *J. Mol. Biol.* **195**, 795–808.
- Saucier, J.-M. & Wang, J. C. (1972) *Nat. New Biol.* **239**, 167–170.
- Kim, R. & Kim, S.-H. (1982) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 451–454.
- Keller, W. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4876–4880.
- Shure, M. & Vinograd, J. (1976) *Cell* **8**, 215–226.
- Depew, R. E. & Wang, J. C. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4275–4279.
- Pulleyblank, D. E., Shure, M., Tang, D., Vinograd, J. & Vosberg, H.-P. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4280–4284.
- Dougherty, G. (1983) *Biosci. Rep.* **3**, 453–460.
- Bowater, R., Aboul-Ela, F. & Lilley, D. M. J. (1992) *Methods Enzymol.* **212**, 105–120.
- Lee, C.-H., Mizusawa, H. & Kakefuda, T. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2838–2842.
- Scatchard, G. (1949) *Ann. N.Y. Acad. Sci.* **51**, 660–672.
- Crothers, D. M. (1968) *Biopolymers* **6**, 575–584.
- Müller, W. & Crothers, D. M. (1968) *J. Mol. Biol.* **35**, 251–290.
- McGhee, J. D. & Von Hippel, P. H. (1974) *J. Mol. Biol.* **86**, 469–489.
- Moore, M. H., Hunter, W. N., d'Estaintot, B. L. & Kennard, O. (1989) *J. Mol. Biol.* **206**, 693–705.
- Frederick, C. A., Williams, L. D., Ughetto, G., van der Marel, G. A., van Boom, J. H., Rich, A. & Wang, A. H.-J. (1990) *Biochemistry* **29**, 2538–2549.
- Chaires, J. B., Fox, K. R., Herrera, J. E., Britt, M. & Waring, M. J. (1987) *Biochemistry* **26**, 8227–8236.
- Roche, C. J., Thomson, J. A. & Crothers, D. M. (1994) *Biochemistry* **33**, 926–935.
- Nicolaou, K. C. & Dai, W.-M. (1991) *Angew. Chem. Int. Ed. Engl.* **30**, 1387–1416.
- Zein, N., Sinha, A. M., McGahren, W. J. & Ellestad, G. A. (1988) *Science* **240**, 1198–1201.
- Zein, N., Poncin, M., Nilakantan, R. & Ellestad, G. A. (1989) *Science* **244**, 697–699.
- Drak, J., Iwasawa, N., Danishefsky, S. & Crothers, D. M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 7464–7468.
- Aiyar, J., Danishefsky, S. J. & Crothers, D. M. (1992) *J. Am. Chem. Soc.* **114**, 7552–7554.
- Hawley, R. C., Kiessling, L. L. & Schreiber, S. L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1105–1109.
- Walker, S., Murnick, J. & Kahne, D. (1993) *J. Am. Chem. Soc.* **115**, 7954–7961.
- Walker, S., Valentine, K. & Kahne, D. (1990) *J. Am. Chem. Soc.* **112**, 6428–6429.
- Uesugi, M. & Sugiura, Y. (1993) *Biochemistry* **32**, 4622–4627.
- Kuduvalli, P. N., Townsend, C. A. & Tullius, T. D. (1995) *Biochemistry* **34**, 3899–3906.
- Mah, A., Price, M. A., Townsend, C. & Tullius, T. D. (1994) *Tetrahedron* **50**, 1361–1378.
- Ikemoto, N., Kumar, R. A., Ling, T., Ellestad, G. A., Danishefsky, S. J. & Patel, D. J. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 10506–10510.
- Depew, K. M., Zeman, S. M., Boyer, S. H., Denhart, D. J., Ikemoto, N., Danishefsky, S. J. & Crothers, D. M. (1996) *Angew. Chem. Int. Ed. Engl.* **35**, 2797–2801.
- Wu, H.-W., Dattagupta, N., Hogan, M. & Crothers, D. M. (1980) *Biochemistry* **19**, 626–634.
- Chaires, J. B., Dattagupta, N. & Crothers, D. M. (1983) *Biochemistry* **22**, 284–292.